AcaStat User Manual

Version 7.2 for Mac and Windows

© Copyright 2011, AcaStat Software.
All rights Reserved.
http://www.acastat.com

Table of Contents

INTRODUCTION 5

GETTING HELP 5
INSTALLATION 5
UNINSTALLING ACASTAT 5
TERMINOLOGY USED IN THE MANUAL 5
ACASTAT STARTUP SCREEN 6

DATA TAB 7

Rows 7
Columns 7
Cells 8
Entering data 8
Editing data 8
Spreadsheet and Data File Limits 8
FILE MENU 9
New 9
Open Data File 9
Save Data File As 9
Data File Information 9
Import Tab Delimited Data 9
Import Comma Delimited Data 9
Tips on Importing Data 9
Pasting Data 10
Importing Text Data Files 10
Tips on using a spreadsheet application to create a csv or txt file 10
Example of Import Error 11
Export Data (Tab Delimited) 11
Export Data (Comma Delimited) 11
DATA MENU 12
Split Large Data Files 12
Merging Data Files to Add Records 12
Create Data Dictionary 13
Use Weight Variable 13
Filter Data 13
Run Procedure 13
EDIT MENU 13
Replace Values 13
Add Row 14
Add Column 14
Delete Row 14
Delete Column 14
Paste 14
Clear 14
VARIABLE MENU 15
Format Variable 15
Variable Name 15

Variable Label 15
Change Decimals 15
Value Labels 15
Missing Values 16
Recoding Values 16
Computing New Variables 17
Combining Two Variables 18
STATISTICAL PROCEDURES 19
CONTROLS 19
Variable List 19
Dependent Variable 19
Independent Variable 19
Control Variable 19
Tips for using the Control Box 20
PROCEDURES 20
Frequencies 20
List Variables 20
Descriptives 20
Explore Means 20
Crosstabulation 20
T-tests of means 21
Analysis of Variance (ANOVA) 21
Pearson Correlation 21
OLS Regression 21
Logistic Regression 21
Diagnostic Accuracy 22
Simulation 22
Random Sample 22
Repeated Sampling 23
Appraisal Statistical Procedure 23
OUTPUT VIEWER TAB 24
FILE MENU 25
New 25
Open Output File 25
Save Output As 25
Print Output 25
EDIT MENU 25
Cut 25
Copy 25
Paste 25
Clear 26
Select All 26
SUMSTATS TAB 27
CALCULATE OPERATIONS 27
STATISTICAL MODULE PULL-DOWN MENU 28
SPREADSHEET 28
SUMMARY OF STATISTICAL PROCEDURES 28
Descriptive Data Module 28

Correlation Data Module 28
Regression Data Module 29
One-Way ANOVA 29
Chi-square 29
Confidence Interval-Means 29
Confidence Interval-Proportions 30
Diagnostic Accuracy 30
Standardized Score 30
T-Test of Means (One Sample) 31
T-Test of Means (Two Sample) 31
Z-test 31
One Sample Proportion 31
Two Sample Proportions 32
SUMSTATS TIPS 32

STATISTICS TUTOR 33

TUTORIAL REVIEW ORDER 33

DECISION TOOLS TAB 34

DECISION TABLES 34
QUEUING THEORY 35
CONSTANT DOLLARS 35
PRICE ELASTICITY OF DEMAND 35

CHARTS TAB 36

SAVE CHART 36
PASTE CHART 37
VARIABLE LIMITS 37
EDITING CHARTS 37

OTHER MENU ITEMS AND ACASTAT FEATURES 38

DATA EXAMPLES MENU 38
DOCUMENTATION 38
HELP MENU 38
Help 38
Statistics Glossary 38
Check for Updates 39
Applied Statistics Handbook 39
Student Workbook 39

REVIEW OF FEATURES 40

DATA SPREADSHEET 40
SUMSTATS 41

PURCHASE INFORMATION 42

Introduction

AcaStat provides students, teachers, researchers, and managers with an inexpensive and easy to use data analysis tool and instructional aid. AcaStat provides a spreadsheet for entering and viewing data, statistical procedures and summary statistics modules, an Output Viewer log, decision tools, and statistics tutorial. A statistics glossary is also provided and purchased versions of AcaStat include the Applied Statistics Handbook.

Getting Help

Application help can be found in this manual and in the Help module provided with AcaStat. A QuickStart guide is also available at http://www.acastat.com/universal.htm.

In addition to application help, statistical help is provided in four formats:

- ✓ A statistics tutorial is linked to the SumStats module.
- ✓ A glossary is provided for statistical terms and definitions.
- ✓ An Applied Statistics Handbook is provided in pdf format for licensed versions.
- ✓ A Student Workbook is provided in pdf format.

Installation

Download AcaStat from http://www.acastat.com. The software is downloaded in a self-extracting file. Save the software to your hard drive and double-click on the file to begin the installation process.

Uninstalling AcaStat

For Mac versions, drag the program file to the trash bin. If you are using a Windows version, use the Windows Control Panel option ADD/REMOVE programs to uninstall AcaStat. Files that you have created or changed must be removed separately by the manual file deletion procedures in Mac and Windows.

Terminology Used in the Manual

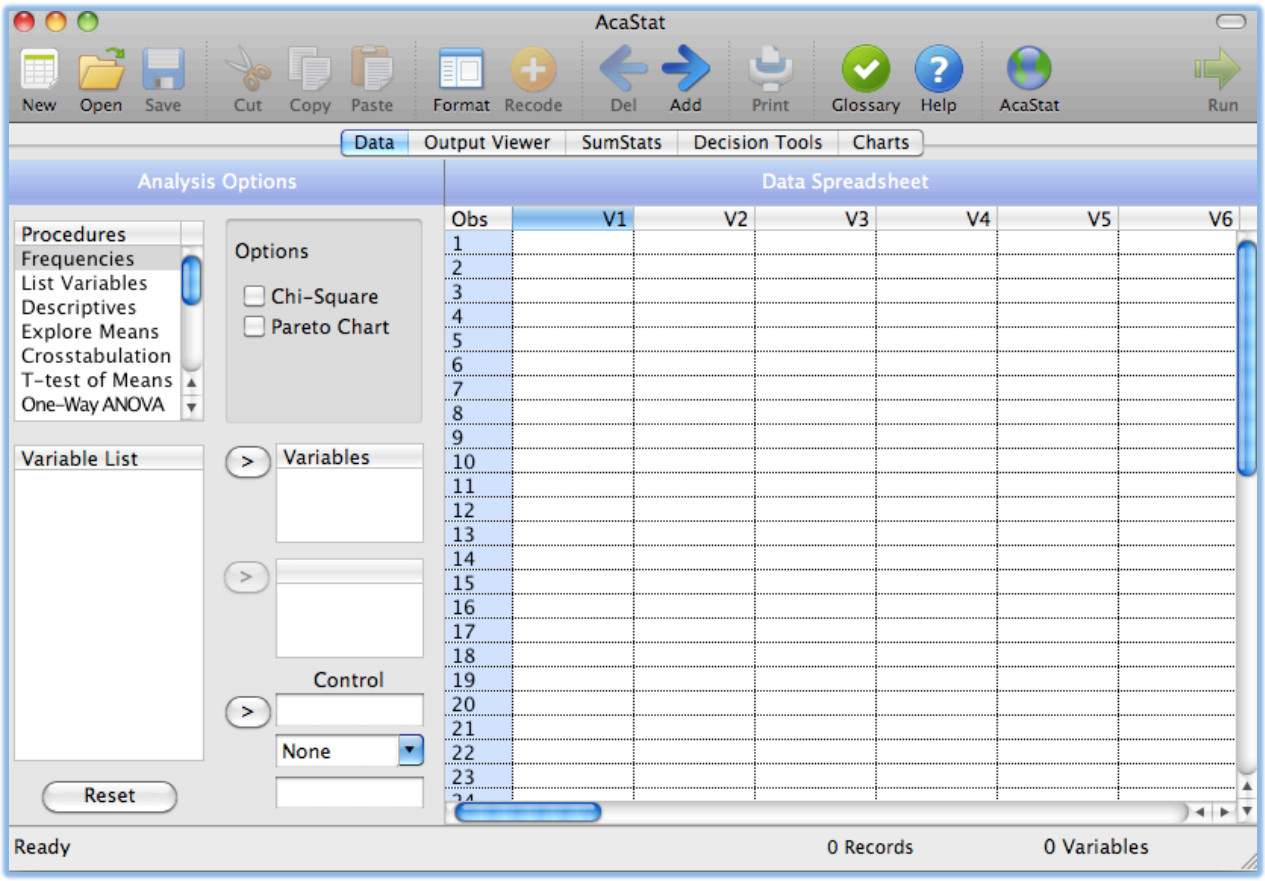
When discussing statistical procedures, the manual will refer to two general types of data. The following terms will be used for convenience and simplicity, but there are variations in these themes that will be apparent to those with statistics training. Please see the Statistics Glossary in AcaStat for more information on these and related terms.

Continuous Data – This refers to data where is reasonable to calculate an average. This type of data is also referred to in statistics as interval/ratio level of measurement or scale data.

Categorical Data – This refers to data where averages would not be meaningful. This type of data is also referred to in statistics as nominal/ordinal level of measurement or discrete data.

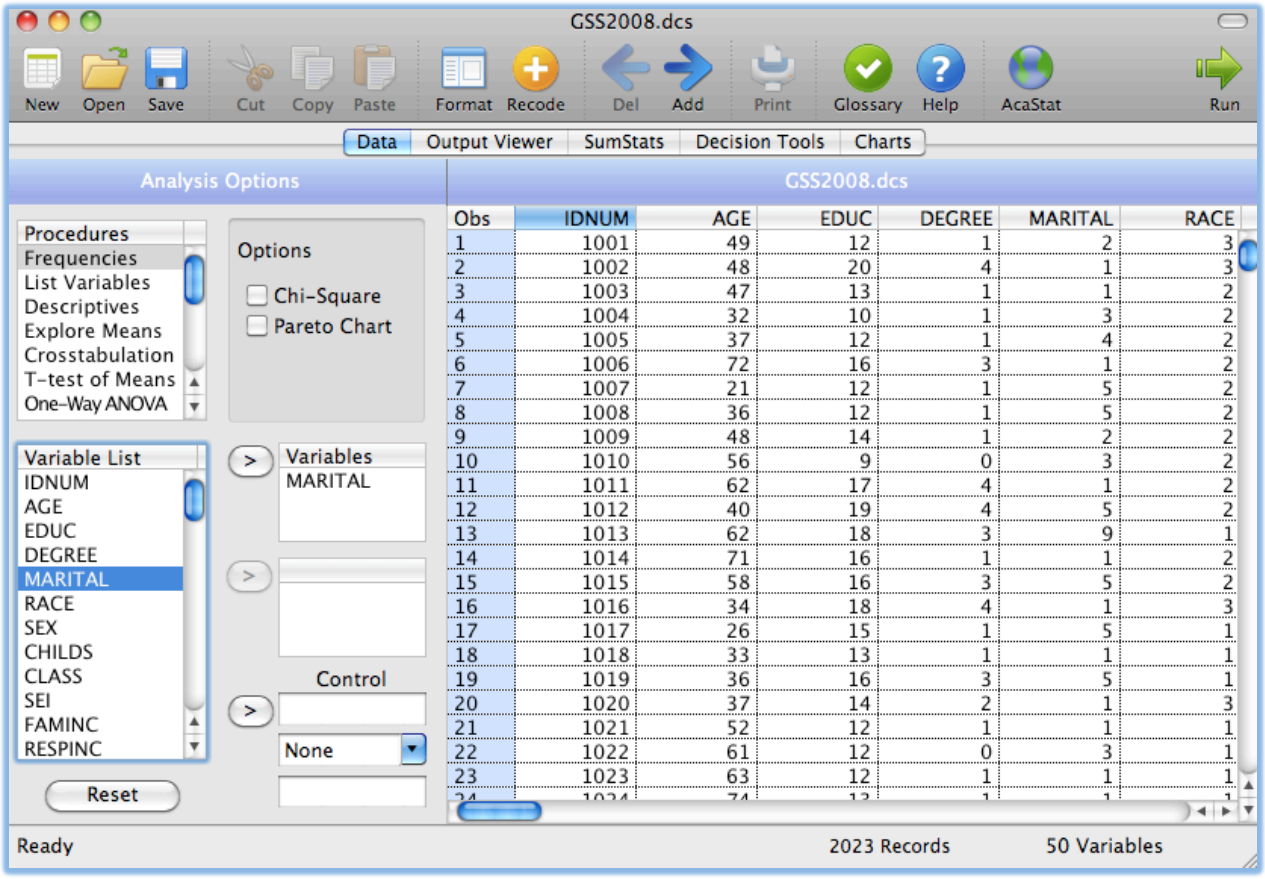
AcaStat Startup Screen

AcaStat starts by displaying the Data tab and an empty spreadsheet. There are also tabs for the Output Viewer, SumStats, Decision Tools, and Charts. A control panel along the right side of the application contains statistical procedure controls. Use the Statistical Procedures panel to select and run analyses. The Output Viewer tab will automatically display the results for the current analysis and will store all analyses conducted after opening AcaStat. The output from all procedures, including SumStats and the tutorial, will be automatically appended to the output log.



Data Tab

The Data tab uses a spreadsheet to create, import, and export data files. Opening or importing a new file will replace the current contents.



Rows

Each row represents one observation (this is also known as the unit of analysis). If the data contain information on individuals, each row should contain the data for one person (observation). If the data contain information on cities, each row should contain the data relevant to one particular city.

Columns

Each column represents data for all records relating to one variable or characteristic. As an example, if a column contains information on personal income in U.S. dollars, all data in that column should represent the income for each person in the data file.

Cells

Each cell represents a value for one observation for one variable. It can contain string data (words) or numerical data (numbers). Most research data files use numerical data to represent words. As an example, for the variable sex, it is common to code Male as "0" and Female as "1". This makes data entry much easier since one number is entered instead of typing a word. It also simplifies recoding variables, identifying controls, and running analyses.

Entering data

To enter data, point and click on a cell and begin typing. Pressing an arrow key will save the entry and move the data entry point to another cell. As an example, pressing the right arrow key will save the entry and move the data entry point to the right. Pressing the Escape (Esc) key will cancel data entry. Pressing the Enter key will move the data entry point down one row.

Editing data

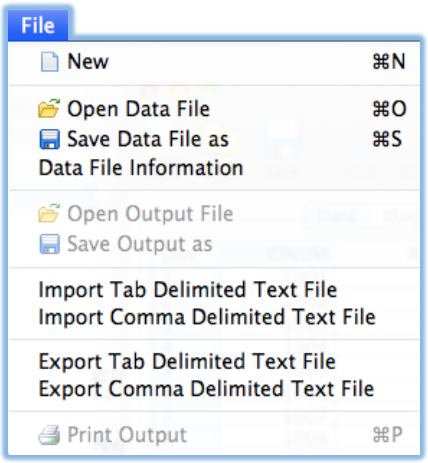
To edit data, click on the cell to edit. Pressing any of the arrow keys will save the entry and move the data entry point to the relevant cell. Do not enter characters such as a Comma, $, or % and avoid the use of strings (words) whenever possible.

Spreadsheet and Data File Limits

The spreadsheet edit capacity is limited to 60 columns. The number of rows is controlled by system memory and speed of the computer. AcaStat was not developed for large files. Files larger than 100,000 observations will slow loading and analyses.

If a file has more than 60 variables, the spreadsheet will display the first 60 variables, although all the variables and observations will be available for analysis. Large files can be reduced using the Split Data module found in the Data pull-down menu. This procedure will create a new file with fewer variables. The new file will have the same number of observations and formatting as the original larger file. If a file has a large number of observations, the random sample statistical procedure can be used to create a more manageable data file.

File Menu



New

Clears all data from the spreadsheet and resets the variable lists.

Open Data File

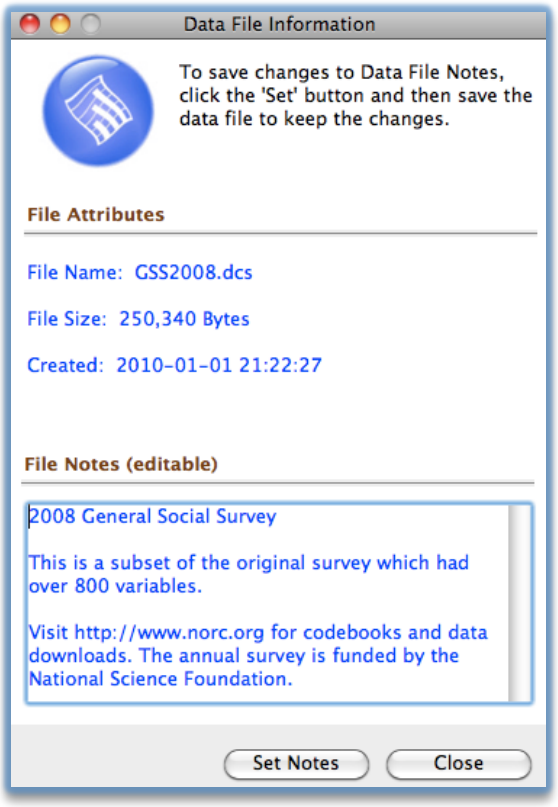
Use to open an existing AcaStat data file. AcaStat data files have the extension "*.dcs". AcaStat files contain data formatting, coding of missing values, and data file notes.

Save Data File As

Saves the contents and any formatting of the data into an AcaStat data file.

Data File Information

Displays file information and also has a section for file notes. Changes to the file notes will not be retained until the data file is saved.


Data File Information
To save changes to Data File Notes, click the 'Set' button and then save the data file to keep the changes.
File Attributes
File Name: GSS2008.dcs
File Size: 250,340 Bytes
Created: 2010-01-01 21:22:27
File Notes (editable)
2008 General Social Survey
This is a subset of the original survey which had over 800 variables.
Visit http://www.norc.org for codebooks and data downloads. The annual survey is funded by the National Science Foundation.
Set Notes
Close

Import Tab Delimited Data

Opens a text file that contains tabs between each variable so AcaStat is able to import the data into a spreadsheet format. These files normally have a filename extension of "*.txt" or "*.dat".

Import Comma Delimited Data

Opens a text file that contains commas between each variable so AcaStat is able to import the data into a spreadsheet format. These files normally have a filename extension of "*.csv". The abbreviation "csv" represents comma separated values.

Tips on Importing Data

Import data from text data files or by pasting data copied from a word processor, spreadsheet, or database table.

Pasting Data

To paste data from a table, select and copy data in a word processor, spreadsheet, or database table. In the Data tab, select "File/Import/Paste Data" or use the paste button on the toolbar. The paste procedure replaces all data in the current spreadsheet. Copied data must be consistent with the tips discussed below.

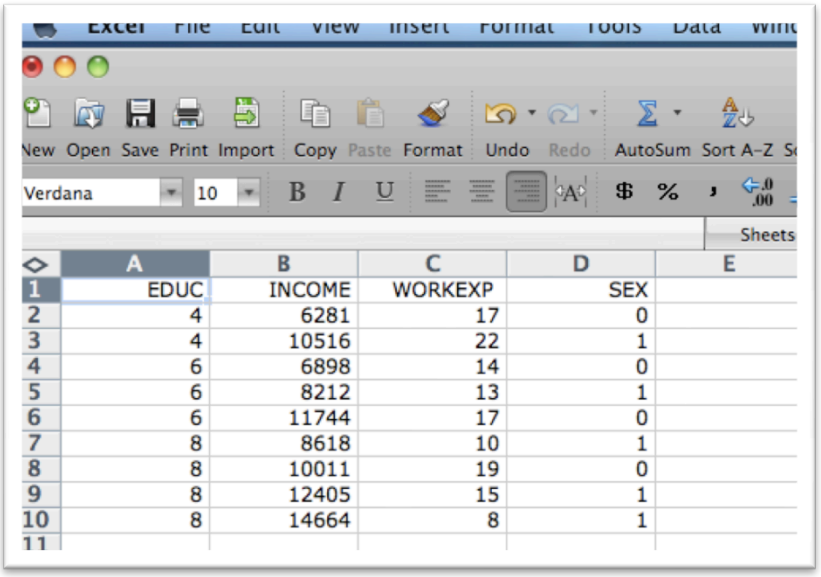
Importing Text Data Files

It is important to note that AcaStat assumes the first row in an imported file contains the variable names. Please see the section below on using a spreadsheet application to create compatible data files.

If the data file is large, it may take a few moments to read the data. Once read, the data will be displayed in the spreadsheet and variable names (column headings) will be visible in the variable list boxes. After importing data, save the file as an AcaStat system file. This will increase loading speed and allows the file to retain variable formatting. As noted earlier, use the Split Data module to reduce the size of very large data files (it is not unusual to see social science data files with over 700 variables).

Tips on using a spreadsheet application to create a csv or txt file

The first row in the spreadsheet should contain a column label that represents the data in that column (keep the labels 8 characters or less). These labels represent the variable names and help identify columns when imported into the AcaStat spreadsheet (see example of spreadsheet format below). To save a delimited text file, use File/Save As and select the file type as either tab delimited or comma separated values (CSV) format.



	A	B	C	D	E
1	EDUC	INCOME	WORKEXP	SEX	
2	4	6281	17	0	
3	4	10516	22	1	
4	6	6898	14	0	
5	6	8212	13	1	
6	6	11744	17	0	
7	8	8618	10	1	
8	8	10011	19	0	
9	8	12405	15	1	
10	8	14664	8	1	
11					

Example of Import Error

When creating a comma separated value file, make sure all commas have been removed from the cell contents before creating the file. If they are not removed, as might be the case with currency data, AcaStat's import procedure will assume the comma is an extra variable. This will cause the cell contents in the AcaStat spreadsheet to be out of alignment with the column. In the example below, the data in rows two and three do not align properly with the appropriate columns.

	AGE	SEX	INCOME	V4
1	65	M	33000	
2		23	M	
3	46000	45	F	

Other import errors may also occur if the text file does not exactly match the format specified above. The following list is not complete, but it contains common errors in importing data.

- No variable names in the first row.
- Blank rows at the top of the data file.
- Special embedded characters in the cells that are neither numbers nor words. This is especially common in data copied from an html table from the Internet.
- Embedded tabs in spreadsheet cells.

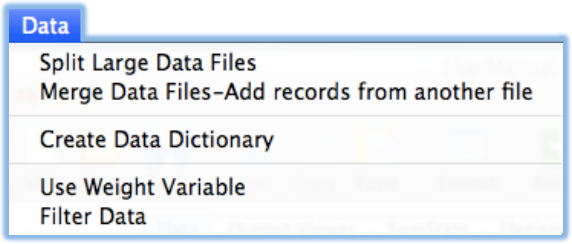
Export Data (Tab Delimited)

Saves the contents of the spreadsheet as a text file that contains tabs between each variable (column). These files are given the filename extension "*.txt".

Export Data (Comma Delimited)

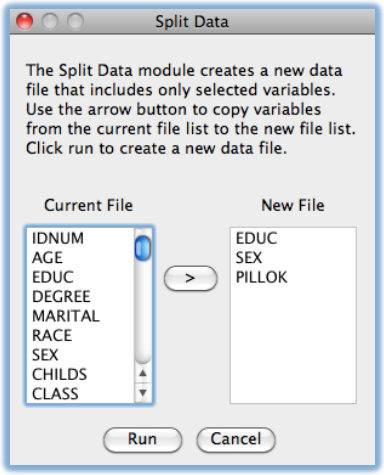
Saves the contents of the spreadsheet as a text file that contains commas between each variable (column). These files are given the filename extension "*.csv".

Data Menu



Split Large Data Files

Use Split Large Data Files to create a new data file that contains a subset of variables from the current file. This feature reduces the size of large data files to increase analysis speed and improve data editing capacity. The new file retains variable formatting.



Click the Edit pull-down menu in the Data tab and select the Split Large Data Files option. A list of the variables in the current data file will be provided. Select variables from the current file list to include in the new file and click the right arrow button. To remove a variable from the new file list, select the variable and click the left arrow button. Click run to create the new data file.

Merging Data Files to Add Records

This procedure can be used to merge two identically formatted AcaStat data files so that the observations in one file are added to the observations of another AcaStat data file. Both files must have identical formatting so that the number of variables is the same in each file and the order of the variables from left to right is identical. This procedure is especially useful if you routinely use the same data formatting (variable names, labels, value labels, missing values) to collect new data. Examples include repeated surveys or monthly data downloads from a website or bank account.

Use the following process:

Create and save a data file with all the formatting set for variable and value labels and missing values.

Create additional AcaStat data files by entering the data or importing a delimited data file. The number and order of the variables (columns) must be identical to the original formatted data file; however, you don't have to enter variable names or labels.

To merge files,

1. Open the file that has the formatting
2. Use the menu item [Data/Merge Data Files] to select and open the second file.

The observations from the second file will be added to the first file.

Create Data Dictionary

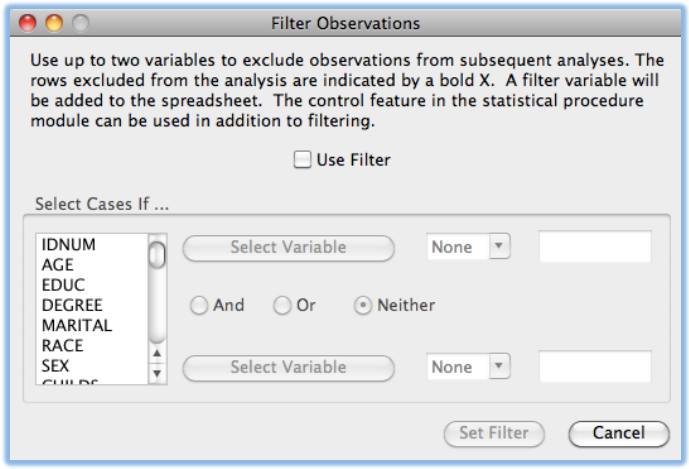
This option creates a record of the data file variables, variable and value labels, and missing values formatting. The record is saved in the Output Viewer for saving or printing.

Use Weight Variable

Select Use Weight Variable to identify a variable for weighting observations. Weights are often used to compensate for over or under sampling and also for producing population values from a random sample. The weights are used in the frequency and descriptive procedures. Weights will not be applied to any other procedures. Only one weight can be selected at a time.

Filter Data

Select Filter Data to view options for excluding observations from the analysis. This module can be used to limit the analysis to a finely defined set of observations. Use up to two variables to exclude observations from subsequent analyses. The rows excluded from the analyses are marked by a bolded X. Resetting the filter removes the filter from future analysis. No data are permanently lost when using a filter.

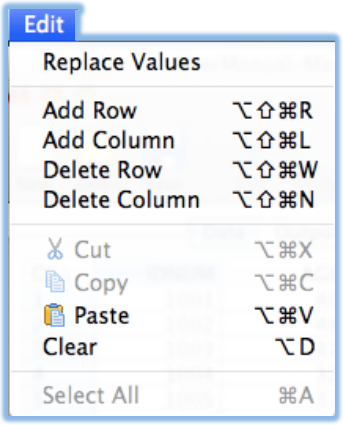


A filter variable will be added to the spreadsheet. If the data file is saved after filtering, this filter variable will be saved and formatted with the filter specifications for later use as a control.

Run Procedure

Select Run Procedure to start an analysis. This has the same function as the Run button on the toolbar.

Edit Menu



Replace Values

Use Replace Values to change cell contents within one variable. This feature is most useful for correcting data entry errors or changing string contents to values (e.g., to change “Male” and “Female” strings so that all “Male” strings are coded as the value 0 and all “Female” strings are coded as a value 1). Use Undo to return the cell contents to the immediate proceeding value.

Procedure

- ✓ Select the variable to search by clicking on the variable name in the listbox.
- ✓ Enter the value or string to search for.
- ✓ Enter the value to use to replace the search parameter.
- ✓ Click the "Replace" button.

Note: It is best to conduct search and replace operations on a duplicate variable so the original variable is not disturbed. Use the Compute procedure to duplicate a variable.

Add Row

Adds one row to the bottom of the spreadsheet to allow additional observations to be added to the data file.

Add Column

This adds one column to the far right of the spreadsheet up to 60 columns. In most cases, this feature will not be used. Columns are added automatically when data are entered into the next to last column.

Delete Row

Deletes the row for the currently selected cell. Delete one row at a time. The change will be complete when the data file is saved.

Delete Column

Deletes the column for the currently selected cell. Delete one column at a time. The change will be complete when the data file is saved.

Paste

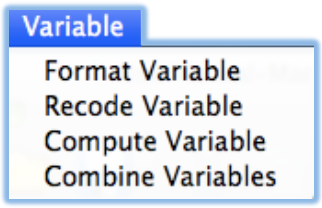
When paste is used in the Data tab, an import procedure places data copied from another application into the spreadsheet. Paste clears the spreadsheet before importing the data. The operation is identical to importing a tab delimited data file.

Clear

If the Data tab is visible, this clears all data from the spreadsheet and resets the variable lists. If an output screen is visible, this will clear all of the contents and the output log.

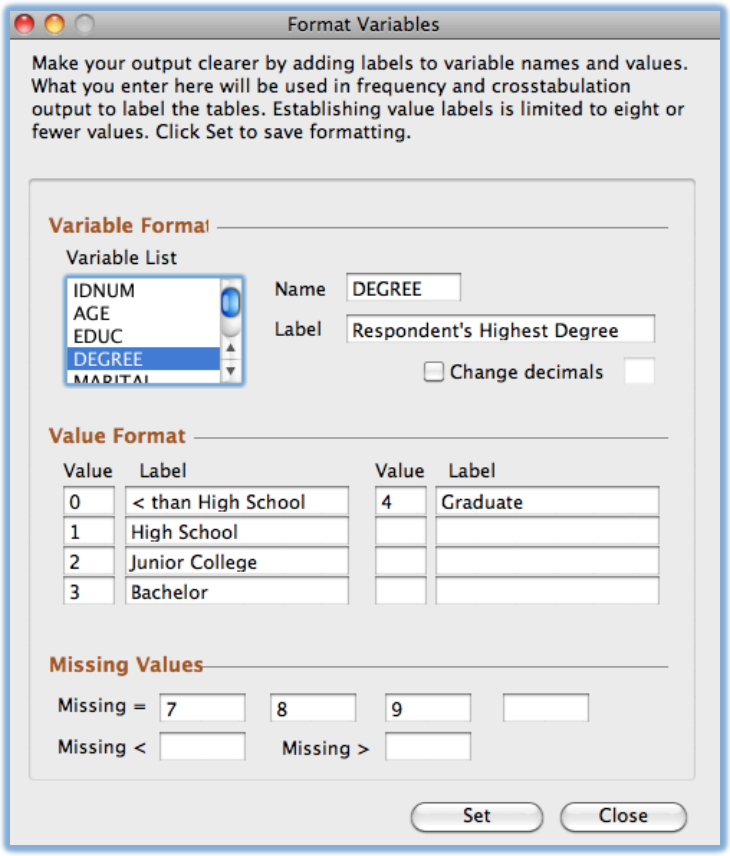
Variable Menu

Format Variable



To make output easier to interpret, create variable names and assign short explanatory labels to variable names and values. Formatting information will be saved with the AcaStat data file and can be changed at any time. If data are exported, the formatting information will not be exported.

Variable Name



Since variable names are used as column labels and are listed in the variable list boxes, it is important to create short but meaningful variable names (8 characters or less). As an example, level of education could be named DEGREE or EDUC.

Variable Label

Use variable labels to create a more detailed description of the variable. The variable label is used to create more meaningful output. Using the above example, DEGREE could be labeled "Respondent's Highest Degree."

Change Decimals

Specify the number of decimals displayed (and stored) for the selected variable. This option will permanently change the variable when the file is saved.

Value Labels

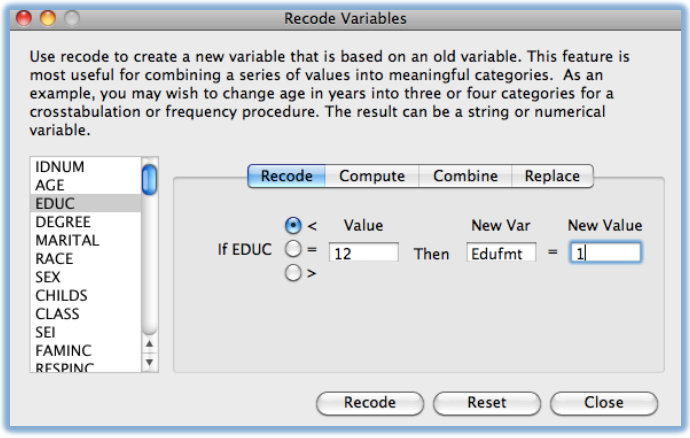
Use value labels to create a more detailed description of the values used in a categorical variable. The value label is used to create more meaningful output. If education is coded 0 through 4 to represent five general categories, the values would be more meaningful if labeled 0="< High School", 1="High school", 2="Junior college", 3="Bachelor", 4="Graduate". Without labels, the output would show the values 0 through 4.

Missing Values

Identify up to four numerical values to be excluded from the statistical analysis (blank cells are always considered missing). Values that are less than (<) or greater than (>) a set value can also be set as missing.

Recoding Values

Use the values of one variable to create a new variable. This is most useful when there are several categories in a variable that can be combined into more meaningful subgroups. An example would be taking years of education and recoding the years into less than high school, high school, some college, and college graduate.



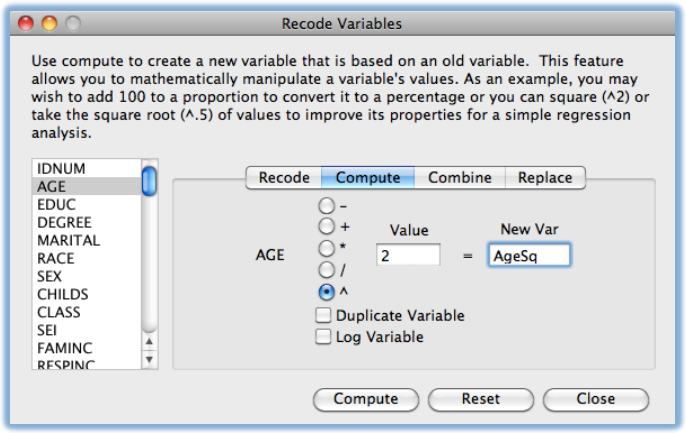
Procedure

- ✓ Select the variable to recode by clicking on the variable name in the listbox.
- ✓ Select the operation to perform by clicking on a radio button. For coding less than 12 years of education, select the "<" button.
- ✓ Enter the value to use as the basis for the recode. For the above example, enter 12.
- ✓ Enter a variable name for the new variable (limit to 8 characters). Example "Edufmt" for education.
- ✓ Enter the new variable's value to represent the old values. For this example, use 1 to represent anyone with less than 12 years of education.
- ✓ Click the Recode button.

A new column will be added to the spreadsheet that has the new variable name and the recoded value. Repeat the process until all values needed for the new variable have been recoded.

Computing New Variables

Create a new variable by mathematically manipulating the values of the original variable. This is most useful when transforming a variable by squaring or taking the square root of the original values. It is also useful for applying a constant to all observations in a variable to remove negative numbers or to develop an index. The Compute procedure can also be used to create an exact duplicate of a variable or create a new variable that represents the natural log of another variable's values.



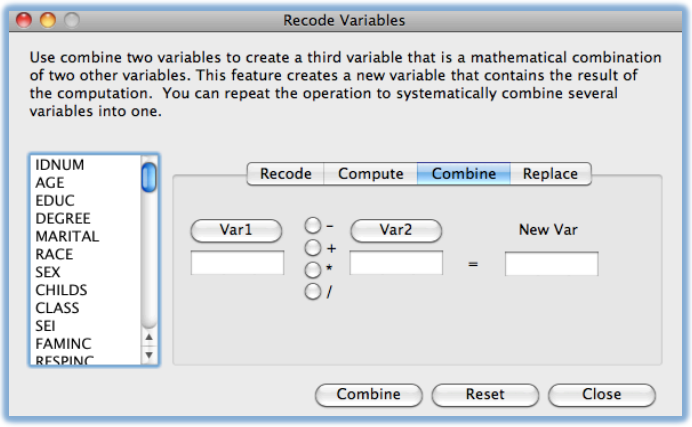
Procedure

- ✓ Select the variable to use for the compute from the listbox.
- ✓ Enter a variable name for the new variable (limit to 8 characters). Example "AgeSq" for squaring age in years.
- ✓ Select the operation to perform by clicking on a radio button.
- ✓ Enter the value to use in the compute statement. Note that ^2=square and ^.5=square root.
- ✓ Click the Compute button.

A new column will be added to the spreadsheet with the variable name and the computed values for all observations.

Combining Two Variables

Use Combine Two Variables to create a third variable that is a mathematical combination of two other variables. This feature creates a new variable that contains the result of the computation. Repeat the operation to systematically combine several variables into one.



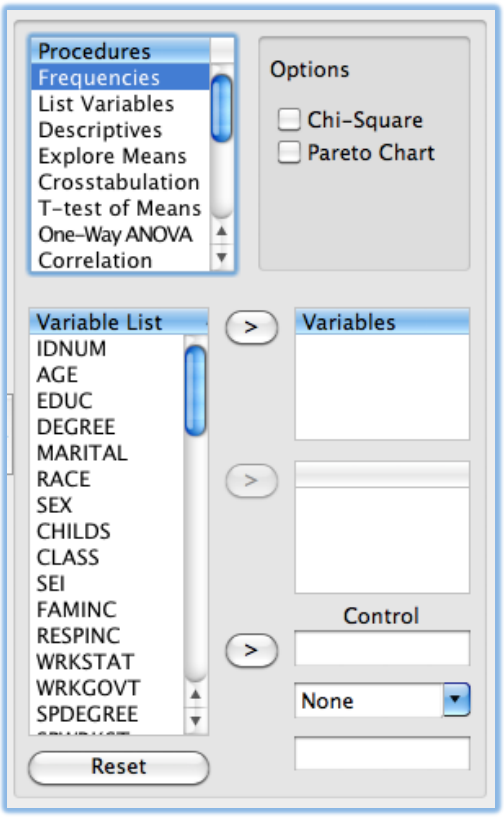
Procedure

- ✓ Select the first variable to use from the listbox and click the Var1 button.
- ✓ Select the second variable from the listbox and click the Var2 button.
- ✓ Enter a variable name for the new variable (limit to 8 characters).
- ✓ Select the operation to perform by clicking on a radio button.
- ✓ Click the Combine button.

Statistical Procedures

The Statistical Procedures control panel is used to select variables and the type of analysis. An options box is displayed for each procedure.

Controls



Variable List

Use the Variable List to select the variables for analysis. As an example, to produce frequencies tables for the variables sex and race, select the sex variable in the variable listbox and click the arrow button to place the variable in the “Variables” listbox. Use the same procedure to select and place the race variable in the listbox. Clicking the Run Procedure toolbar button will produce frequencies for both variables. Output from the analyses will be displayed automatically and recorded in the Output Viewer tab.

Dependent Variable

The topmost variable listbox is used for all univariate analyses (Frequencies and Descriptives) and one bivariate analysis (Correlation). At least one analysis or dependent variable must be placed into the Variables listbox.

Independent Variable

With the exception of Frequencies, Descriptives, and Correlation procedures, an independent variable must also be selected (below the Variables listbox). In the case of t-tests, subgroup values must be set in the options box for the independent variable so AcaStat compares the correct groups.

Control Variable

A control can be established for all procedures. This limits the analysis to a specific subgroup. As an example, if interested in the correlation between income and education for females only, locate the Sex variable name from the variable listbox, click the arrow button to place the variable into the control box, and click the "=" button and indicate the value used for females. This would result in a correlation procedure that excludes all males. The control should be a value (number) not a string (character).

The "Missing" option limits the analysis to records that have missing data for the control variable.

Tips for using the Control Box

If a control variable is selected without establishing a control value, the statistical procedure will run a separate analysis for each value of the control. This is handy for categorical controls such as sex and race but is not recommend for continuous controls such as age, body weight, and income unless these are first recoded into subgroups as categorical values.

Note: Please remember to remove the control variable before running another procedure.

Procedures

One of the statistical procedures must be selected to run an analysis.

Frequencies

This produces a listing of all the values in a variable, the number of times the values occur in the variable, and the percent of cases represented by each value. Frequencies can be produced on numerical and string data.

List Variables

This produces a listing of observations for up to ten variables. You can control the number of observations listed or allow all observations for the selected variables to be listed.

Descriptives

Calculates measures of central tendency (mean, median) and variation (variance, standard deviation). Descriptives requires continuous data.

Explore Means

Explore descriptive statistics (sum, count, mean, sample standard deviation, standard error, 95% confidence interval) for each subgroup in the selected categorical variable(s). Explore requires continuous data for the analysis variable. The categorical variable can be a numeric or string variable.

Crosstabulation

This is a very useful procedure for representing the association between two categorical variables using a contingency table. Each cell in a contingency table represents a subgroup. For each cell, the count, row percent, column percent, and total percent are reported. This procedure also automatically produces a chi-square statistic. Crosstabulations are easiest to interpret when the number of rows and columns in the contingency table are 5 or less.

T-tests of means

Used to test for the difference between two sample means. For t-tests, the dependent variable must be a continuous variable to produce the means for analysis. The independent variable must be a categorical variable (often it will be a dichotomous variable such as sex: 0=male, 1=female). Values representing the comparison groups for the independent variable must be set in the option frame. These values are used to separate the dependent variable into two groups. Be careful to use the correct values or there may be an error reported that indicates there are not enough cases to compute the t-test. If there are more than two values in the independent variable, the other subgroups will be excluded from the analysis.

Analysis of Variance (ANOVA)

Compare the difference between two or more sample means. Like t-tests, the dependent variable must be a continuous variable to produce the means for analysis. The independent variable must be a nominal level variable (e.g.: race: 1=white, 2=black, 3=Asian, 4=Hispanic. or ordinal (political ideology: 1=conservative, 2= moderate, 3=liberal). ANOVA optional output includes summary statistics for each subgroup and the Bonferroni post hoc test for multiple comparisons.

Pearson Correlation

Calculates the Pearson correlation coefficient to represent the association between two continuous variables. There are two options: create matrix or a scatterplot. The scatterplot option provides basic statistics and displays a visual representation of the association between two variables. For multiple comparisons, a correlation matrix is often the best choice.

OLS Regression

Calculates simple regression (one independent variable) or multiple regression (two or more independent variables). Ordinary least squares (OLS) regression requires a continuous dependent variable. As an option, ordinary least squares regression will add new variables to the data representing residuals or predicted values for each observation. Summary statistics for the observations included in the model are automatically included in the output.

Logistic Regression

Logistic regression requires a dichotomous (two values) dependent variable. Logistic regression will automatically produce summary statistics and, as an option, will add a new variable to the data representing the predicted values for each observation.

Diagnostic Accuracy

Diagnostic accuracy compares a test result to a dichotomous reference standard. It is used to evaluate the sensitivity and specificity of a diagnostic test. The reference variable can only be two values (usually 0 for non-disease and 1 for disease). This represents the gold standard to compare to the test results.

Each cell in the table holds counts. The test variable can be nominal, ordinal, or interval/ratio. The lowest value for the reference standard must always represent non-disease. The cutpoint (T+ >) represents the point beyond which the test is positive (disease indicated).

As an example, if possible test results range from 1 to 4, a cutpoint of 1 means a negative test result and a positive test is indicated by a value of 2, 3, or 4. You should carefully evaluate the sensitivity and specificity for one cutpoint to ensure you have configured the data properly before interpreting the results.

Simulation

Two simulation procedures can be used for manipulating data in the spreadsheet. They were designed to help students develop a better understanding of random sampling. The Random Sample option creates a random sample from a larger data file. Assuming the larger data file represents the entire population, a smaller random helps students discover how summary sample measures such as means and proportions will vary from population parameters.

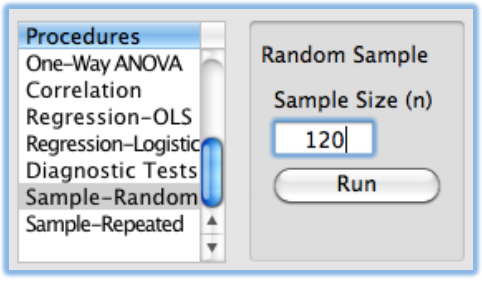
The Repeated Sampling procedure conducts multiple random samples to build a data set that reflects summary measures for each random sample. This procedure helps students test the Central Limit Theorem:

Central Limit Theorem

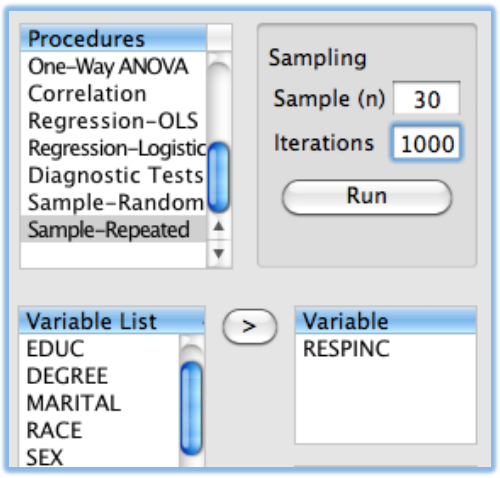
As sample size increases, the sampling distribution of means approximates a normal distribution and is close to normal at a sample size of 30.

Random Sample

This procedure is used to create a data file that is a randomly selected subset of the current spreadsheet. Enter the sample size desired. Please note that the sample size must be less than total number of observations in the current data file. Click the Sample button to replace the current data with a random sample.



Repeated Sampling



This procedure creates repeated random samples for one variable in the current spreadsheet. For each sample (iteration), the mean and standard deviation are computed and placed into the spreadsheet. This results in a new data file where each row represents the summary statistics for one random sample.

Enter the size of the samples (n) and the number of random samples (Iterations) desired. Select a variable to conduct the random sampling on and use the arrow button to place it in the list box. Only one variable will be used for the random sampling. Please note that the sample size must be less than the total number of observations in the current data. Click the Run button to replace the current data with a sampling distribution. When the sampling is complete, AcaStat will run the Descriptives procedure on the mean variable of the sampling data file.

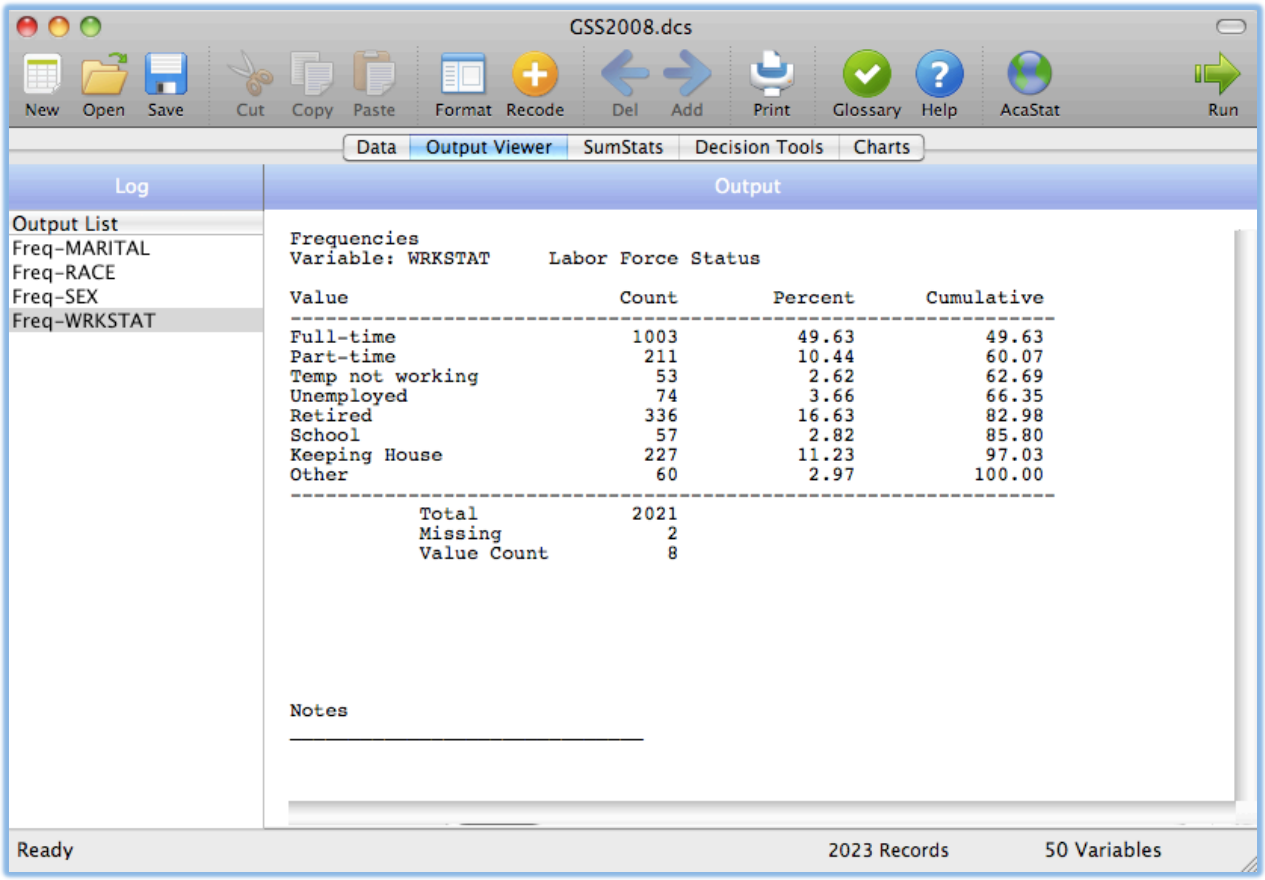
Appraisal Statistical Procedure

The Appraisal procedure is used to determine tax assessment values of real estate. Use of this procedure requires data on the market (selling) price and the assessed value of homes to compute ratios and other statistics. A practice data file “Tax Assessments” is included with AcaStat.

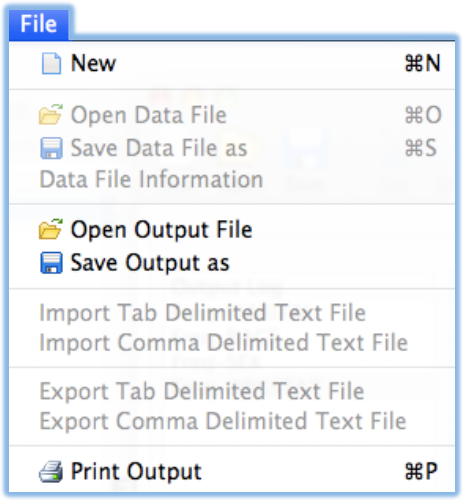
Output Viewer Tab

The Output Viewer tab displays the results of the analysis in a text window for editing, printing,and saving as a text file. Each time a statistical procedure is run, the results are automatically sent to the output log. Notes can be added to the output before printing or saving.

The Output Viewer uses Courier font. This is a fixed font that ensures proper alignment of the tables. When copying output to a word processor, the default font may not be Courier. This will result in output that is out of alignment. To correct this, highlight the output in the word processor and change the font to Courier. The font size may also need to be reduced to ensure the output fits widthwise on one page.



File Menu



New

Click the "New" button to erase all output.

Open Output File

This feature will load output saved from AcaStat statistical software products to include Windows AcaStat Plus and AcaStat Universal for Mac and Windows.

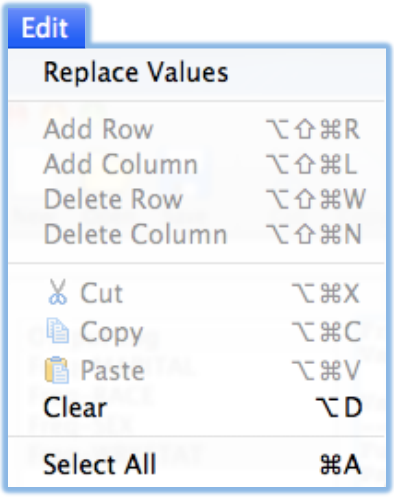
Save Output As

This menu item saves output as a text file with the file extension “*.txt”. This file can be opened in any word or spreadsheet application. Save all output or selected items in the output log. To select multiple items in the Mac version, press and hold the Command key while selecting items in the output log. For the Windows version, use the Control key for the same function.

Print Output

This menu item prints all output or selected items in the output log. To select multiple items in the Mac version, press and hold the Command key while selecting items in the output log. For the Windows version, use the Control key for the same function. By default, AcaStat prints a date and time header and a Notes section on each page printed.

Edit Menu



Cut

Use this menu item to remove selected output text and save the text to the system clipboard.

Copy

Use this menu item to copy selected output text to the system clipboard.

Paste

Use this menu item to paste text from the system clipboard into the output screen.

Clear

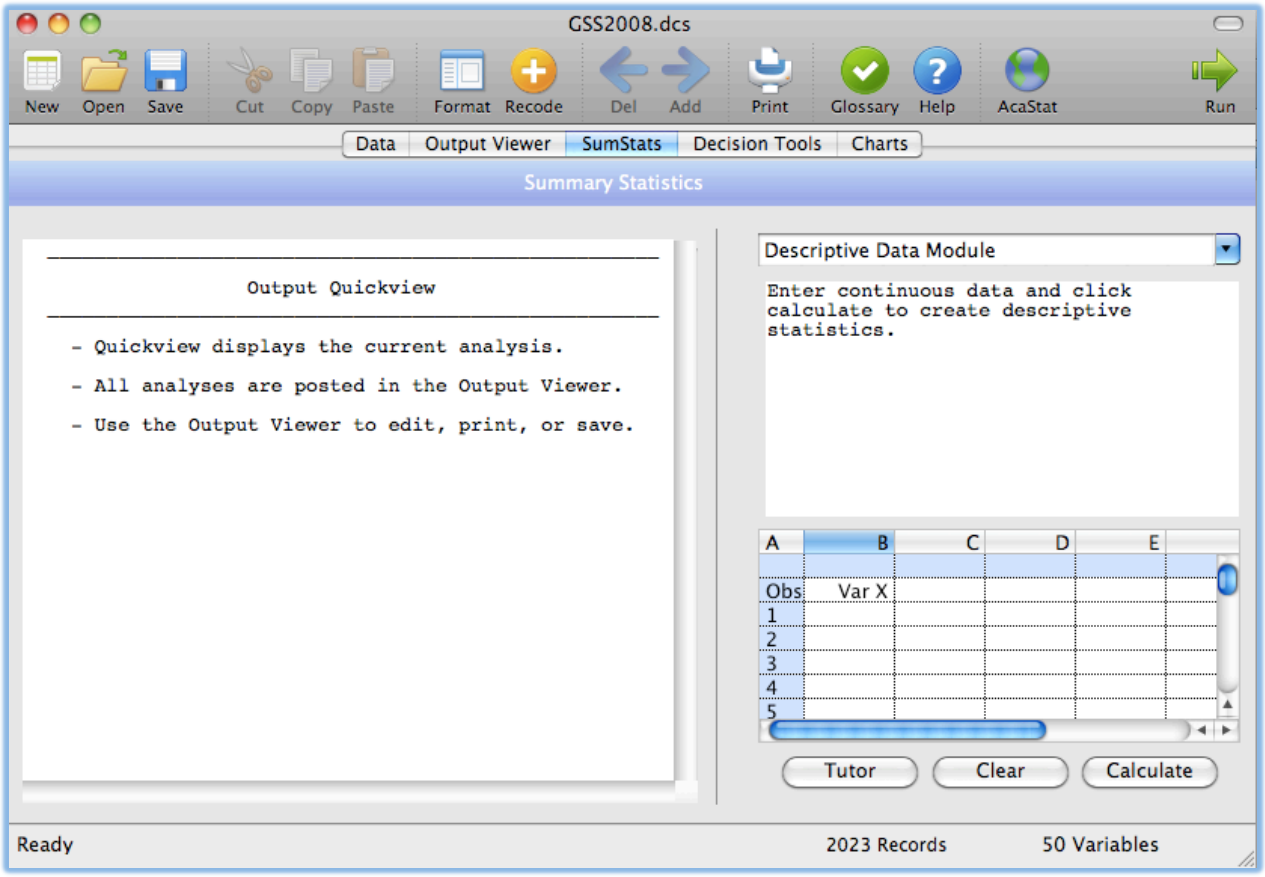
Use this menu item to delete all output log entries.

Select All

Selects all the text in the output screen.

SumStats Tab

SumStats analyzes summary statistics and serves as a resource for the Statistics Tutor. Without access to the raw data, it is difficult to conduct statistical comparisons of means, counts, and proportions without using hand calculations and statistical tables. SumStats solves this problem by allowing the use of summary data to calculate significance tests and other common descriptive calculations.



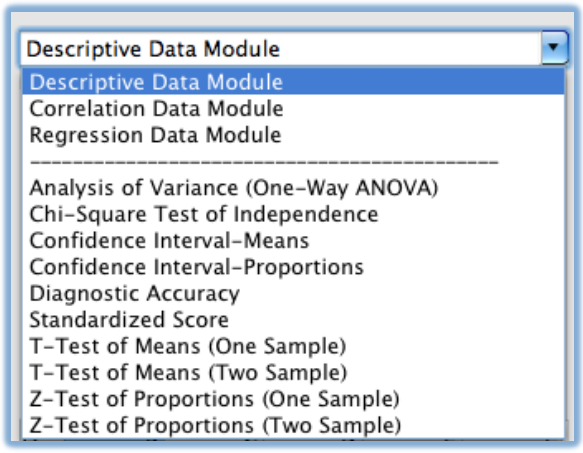
Most statistics courses use summary statistics to teach the fundamentals. SumStats is very useful for checking the results of hand calculations.

Calculate Operations

When the required data are entered, press the Calculate button to run the analysis. If the Tutor is not in operation, the output from the statistical module is displayed in the text box. If the Tutor is in use, the output is placed in the SumStat information box under the statistical module drop-down listbox. The output displayed during the tutorial will be modified to fit in the information box and may contain fewer items than normally included in the procedure results. All calculations using SumStats also post a copy of the standard output in the Output Viewer tab.

Statistical Module Pull-down Menu

SumStats offers a wide range of statistical modules from descriptive statistics to regression. The three modules listed at the top of the pull-down listbox above the dash line replicate raw data analysis normally conducted via a file in the Data tab. The modules below the dash line use summary statistics such as counts, proportions, and means.



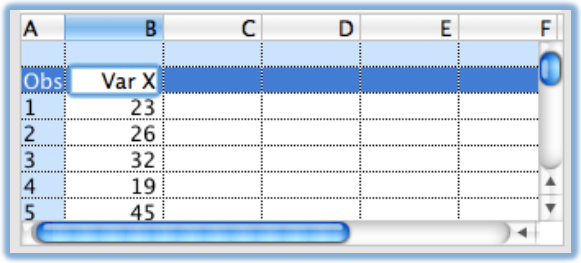
Spreadsheet

Move from one cell to another in the SumStat spreadsheet using keyboard arrow keys or the Enter key. Only numerical data should be entered and it must not contain commas or special characters such as % or $. Decimal points are allowed. Column labels (Var X, etc.) can be edited but limit the label to eight characters or less.

Summary of Statistical Procedures

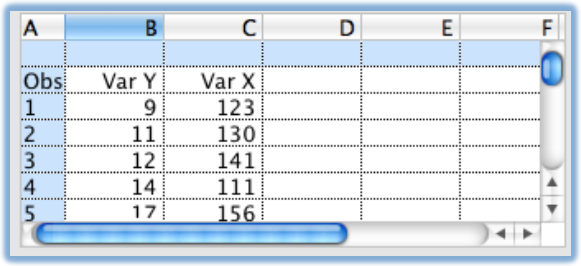
Descriptive Data Module

Enter up to 50 observations in the column labeled Var X. Only data entered under the Var X column are used in the analysis. Click the Calculate button to create summary descriptive statistics such as mean, median, variance, and standard deviation.



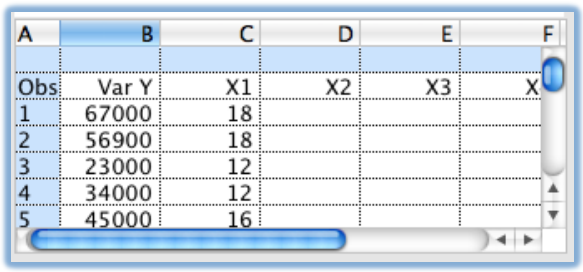
Correlation Data Module

Enter up to 50 observations in the columns labeled Var X and Var Y. Only data entered under the Var X and Var Y columns are used in the analysis. Click the Calculate button to create the Pearson Correlation Coefficient and a scatterplot. Correlation data requires a minimum of 5 observations.



Regression Data Module

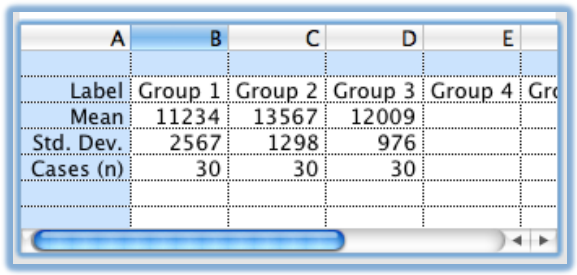
Enter up to 50 observations in the columns labeled Var Y and X1 through X5. Only data entered in these columns are used in the analysis. At a minimum, there must be entries for Var Y (the dependent or response variable of the regression model) and X1 (the independent variable). Click the calculate button to create the OLS regression output. This module can run simple and multiple regression models.



	A	B	C	D	E	F
Obs		Var Y	X1	X2	X3	X
1		67000	18			
2		56900	18			
3		23000	12			
4		34000	12			
5		45000	16			

One-Way ANOVA

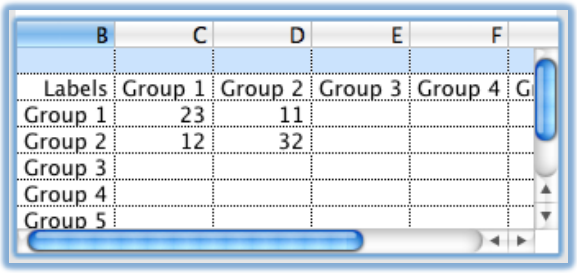
Each column represents one sample distribution. At least two columns (the two left columns) must have data for SumStats to compute statistics. For each sample group, enter the mean, standard deviation, and number of cases (n). Once the data are entered as indicated, click the calculate button.



	A	B	C	D	E
Label		Group 1	Group 2	Group 3	Group 4
Mean		11234	13567	12009	
Std. Dev.		2567	1298	976	
Cases (n)		30	30	30	

Chi-square

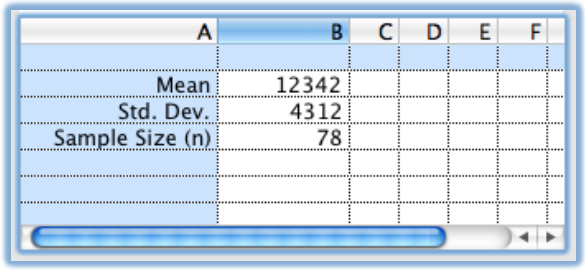
Enter frequencies (counts) in the boxes and click the calculate button. The columns represent one variable with up to six subgroups (e.g., Protestant, Catholic, and Muslim are subcategories for the variable "Religion"). The rows represent a second variable with up to six subgroups. Start with the left column. At least two columns (the two left columns) and two rows (the top two rows) must have data for SumStats to compute statistics.



	B	C	D	E	F
Labels	Group 1	Group 2	Group 3	Group 4	G
Group 1	23	11			
Group 2	12	32			
Group 3					
Group 4					
Group 5					

Confidence Interval-Means

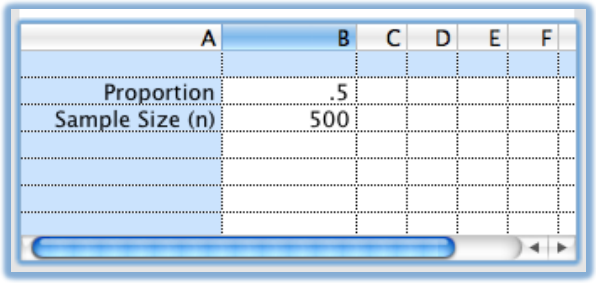
Enter the mean, standard deviation, and sample size and click the calculate button to estimate the margin of error. SumStats will produce 90%, 95%, and 99% confidence intervals (in order from top to bottom).



	A	B	C	D	E	F
Mean		12342				
Std. Dev.		4312				
Sample Size (n)		78				

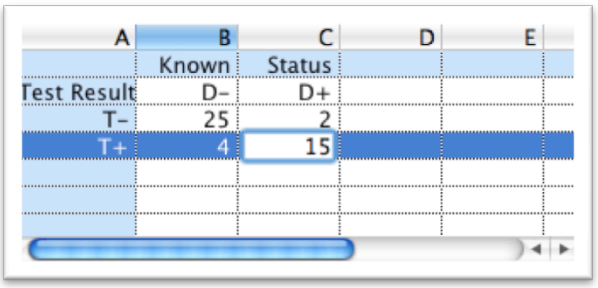
Confidence Interval-Proportions

Enter the proportion and sample size and click the calculate button to estimate the margin of error. SumStats will produce 90%, 95%, and 99% confidence intervals (in order from top to bottom). Convert percentages into proportions (e.g. 25% = .25) before entering them into this module.



Diagnostic Accuracy

This procedure is used to evaluate the accuracy of a diagnostic test in identifying disease presence or absence. The 2x2 grid uses counts from a crosstabulation between known disease status and the diagnostic test results for each patient. To use this module, enter the counts in the relevant boxes and click the calculate button.

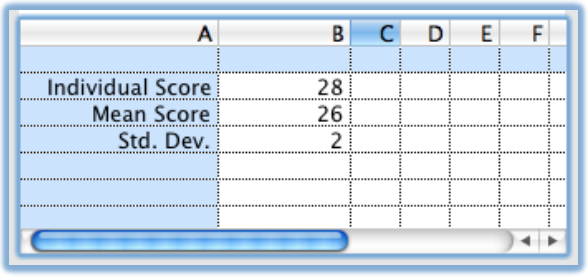


If the counts are not already provided, counts can be produced using the following steps:

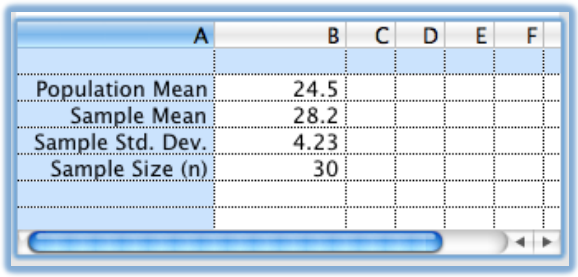
a) create a data file in the Data tab where the rows represent patients and there is one column for known disease status for each patient (0=No Disease, 1=Disease) and one column for test results (0=Neg, 1=Pos);
b) run the crosstabulation procedure where the test variable is the row variable and the disease status is the column variable; and,
c) enter the counts from the crosstabulation into SumStats.

Standardized Score

To use the standardized score module, enter one score from a distribution of scores in the first box. Enter the mean of the distribution in the second box. Enter the standard deviation of the distribution in the third box. Click the calculate button to compute the z-score.



T-Test of Means (One Sample)

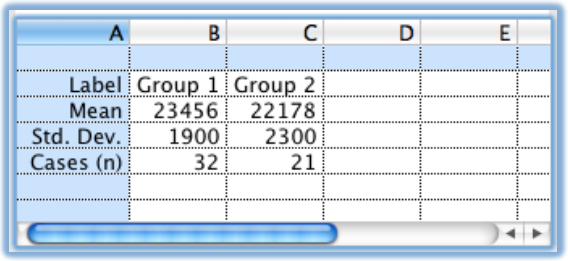


	A	B	C	D	E	F
	Population Mean	24.5				
	Sample Mean	28.2				
	Sample Std. Dev.	4.23				
	Sample Size (n)	30				

To conduct a t-test module for one sample mean, enter the data in the boxes as indicated and click the calculate button. The population mean is the hypothetical mean that is compared to the sample mean. If left blank, SumStats will assume it is 0.

This module may also be used for paired t-tests. For a paired t-test, compute the difference for each case between Var1 and Var2 and calculate the mean and standard deviation for this difference. Enter this information into the relevant Sample boxes and enter 0 for the population mean. This will test the null hypothesis that the paired difference between the means of Var1 and Var2 is not significantly different from zero.

T-Test of Means (Two Sample)



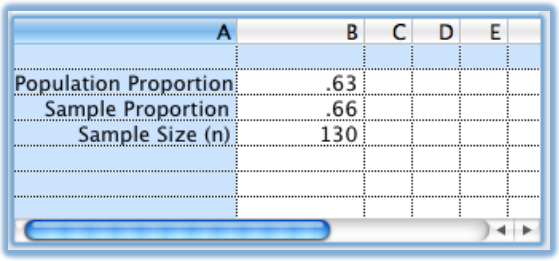
	A	B	C	D	E
	Label	Group 1	Group 2		
	Mean	23456	22178		
	Std. Dev.	1900	2300		
	Cases (n)	32	21		

To conduct a t-test module for two sample means, enter the data in the boxes as indicated and click the calculate button. An F-test for homogeneity of variance is provided along with t-test p-values for samples with equal and unequal variance.

Z-test

There are two z-test procedures that test the difference between two proportions. A common error is to fail to convert percentages to proportions (e.g. 25% = .25), so please remember to convert percentages before entering them into this module.

One Sample Proportion

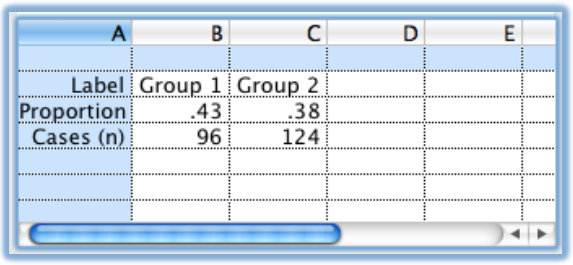


	A	B	C	D	E
	Population Proportion	.63			
	Sample Proportion	.66			
	Sample Size (n)	130			

To conduct a Z-test for one sample proportion, enter the data in the boxes as indicated and click the calculate button. The population proportion is the hypothetical proportion that is used to compare to the sample proportion. If left blank, SumStats will assume it is 0.

Two Sample Proportions

To conduct a Z-test for two sample proportions, enter the data in the boxes as indicated and click the calculate button. It is important to remember that, when comparing percentages, first convert them to proportions (e.g., 75% = .75).



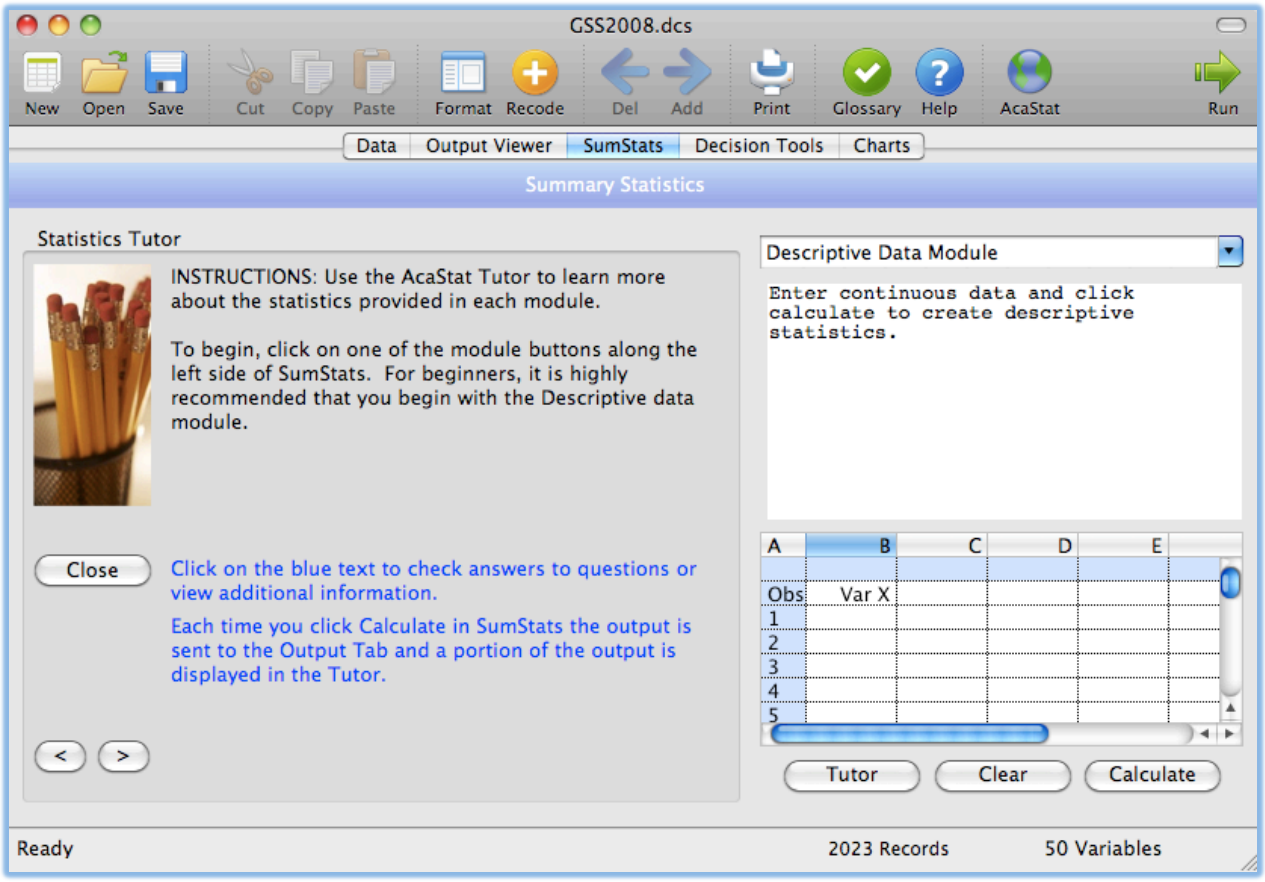
	A	B	C	D	E
	Label	Group 1	Group 2		
	Proportion	.43	.38		
	Cases (n)	96	124		

SumStats Tips

- Module specific instructions are presented when a module is selected.
- When the Tutor is in use, instructions are replaced with summarized results from a calculation.
- Use the arrow keys to move between boxes in the spreadsheet.
- Analyses are automatically added to the Output Viewer tab.
- Entering a label is not required but is useful for interpreting output.

Statistics Tutor

A Statistics Tutor is provided for a review of basic statistics. After starting the Tutor by clicking the Tutor button, select the SumStat module of interest. The Tutor will automatically display the tutorial for that particular module. When using the tutorial, the SumStat output will be displayed in the SumStat information screen. More detailed output is also available in the Output Viewer tab. The Tutor displays black instructional text and blue questions and additional information. Click on the blue text to check your answers to questions or for more information regarding the subject being covered. The licensed version of AcaStat enables the Tutor for all modules. The evaluation version limits the tutorial to the first lesson (descriptive statistics).



Tutorial Review Order

If you are new to statistics or just need a quick refresher, begin with descriptive statistics followed by standardized scores, confidence intervals, z-tests, and t-tests. The tutorial cannot replace a comprehensive course on statistics but it can be used to review and practice statistical concepts and techniques.

Decision Tools Tab

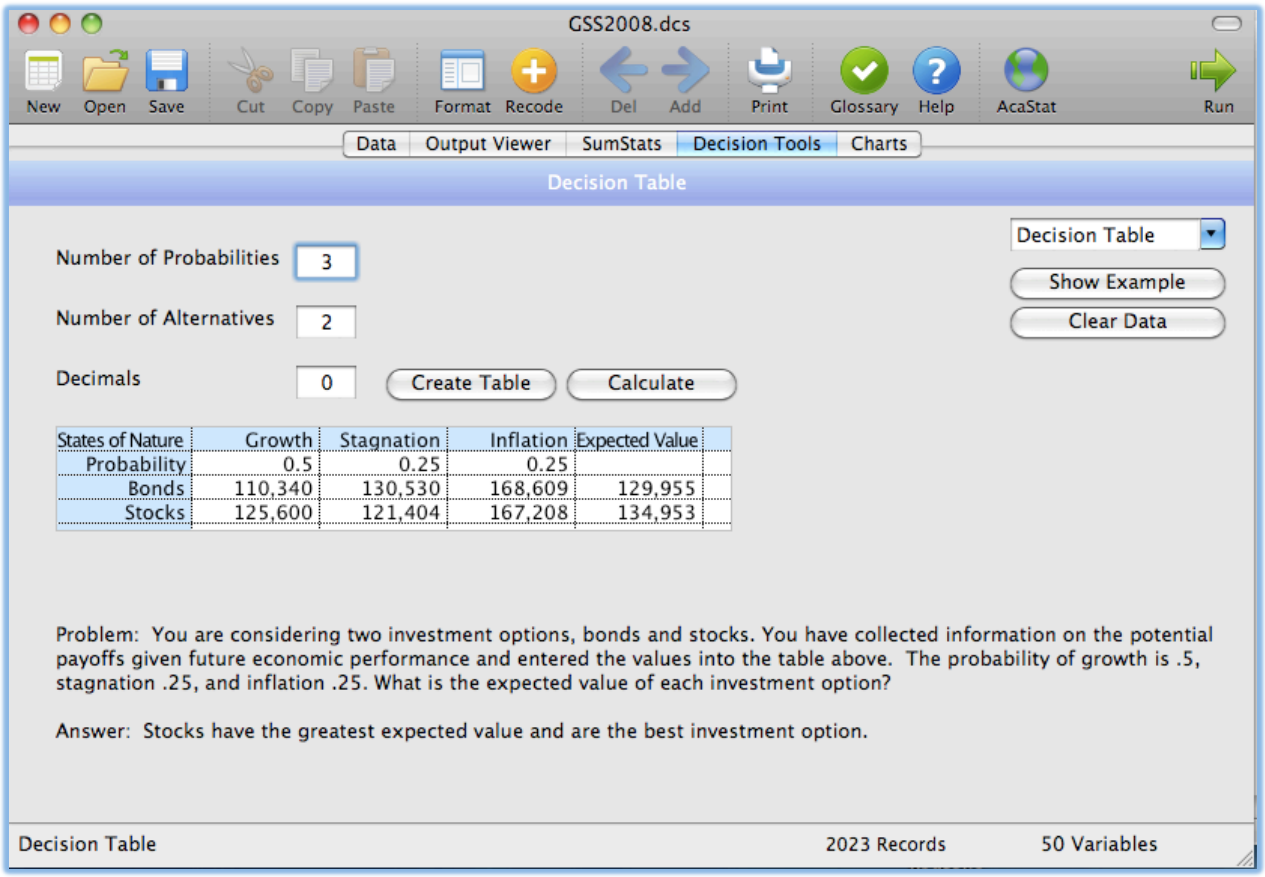
The Decision Tools tab provides four tools for calculating constant dollars and experimenting with decision tables, queuing theory (waiting lines), and price elasticity of demand. An example is provided for each tool. The results are displayed in the Decision Tool tab and are added to the output log in the Output Viewer tab. A good reference text for some of these tools is *Fundamentals of Management Science*, by Efraim Turban & Jack R. Meredith.

Decision Tables

Decision tables quantify in table form the impact of decision options under uncertainty. Tables will typically contain the following elements:

1. Alternative courses of action
2. States of nature
3. Probabilities of states of nature (must sum to 1)
4. The payoffs

An example is provided for each decision tool. The example for the Decision Table is displayed below.



Queuing Theory

Queuing theory is Used to evaluate the performance of service systems where waiting lines may form and customers arrive at random times (i.e., license bureau, tax assessors office, emergency room). To evaluate a waiting line, you need to know the average arrival rate, average service rate, and the number of servicers.

Constant Dollars

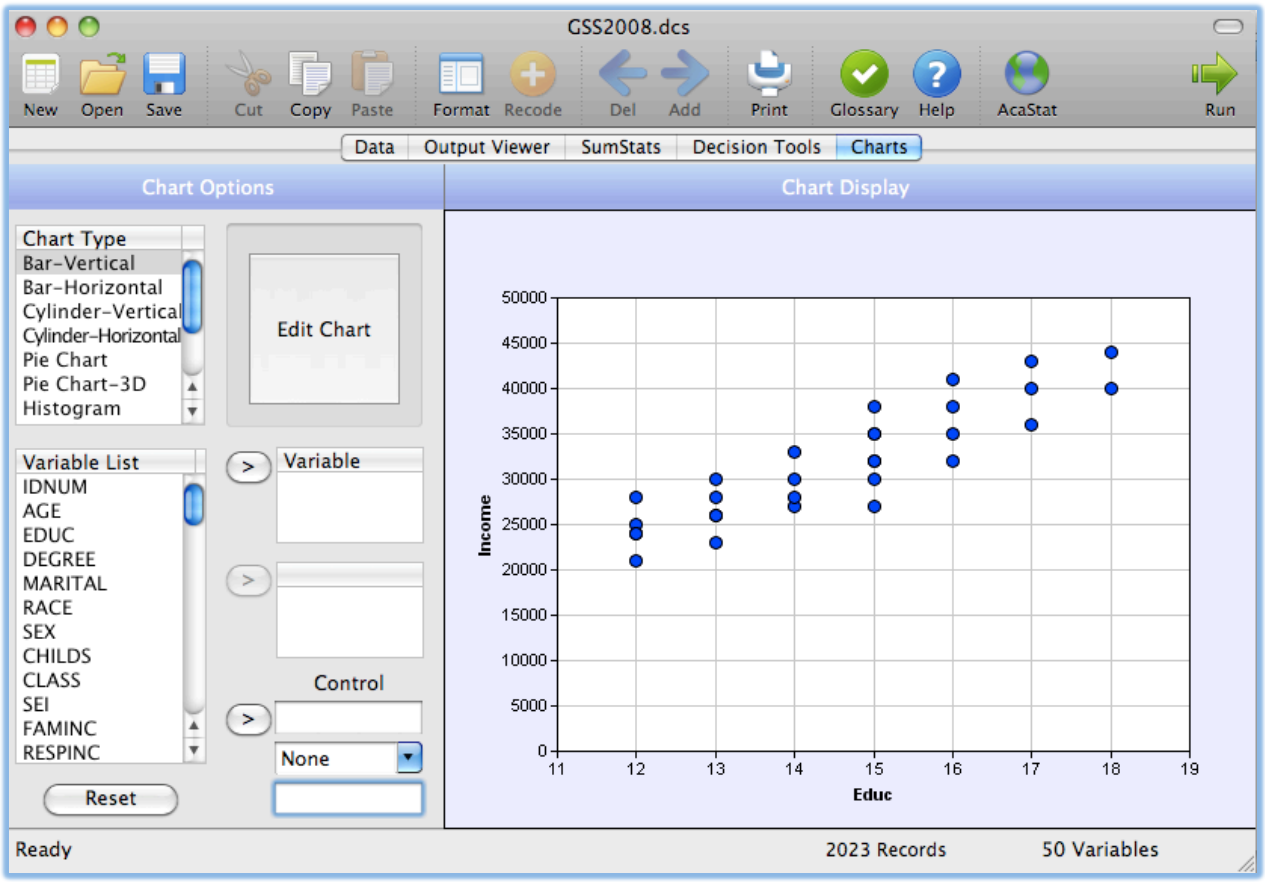
This tool uses the universal consumer price index (CPI-U) for the United States to adjust for inflation. The module will create and apply an inflation and deflation estimate. If you prefer to use another price index, override the default CPI by entering * followed by a CPI you desire.

Price Elasticity of Demand

This tool uses the change in demand that occurs when a price is changed to estimate price elasticity. Given the estimate of elasticity, the tool will attempt to find a breakeven price through iterations of price. You can include both fixed and variable costs.

Charts Tab

The Charts tab produces graphic displays of data. Charts can be saved, printed, and copied by using the toolbar buttons or menu pull-down. Charts can also be formatted to adjust background, titles, and decimals. The control variable option available for conducting statistical procedures is also available to chart subgroups of data.



Save Chart

If using Windows and your system does not have QuickTime installed, you will only be able to save as a *.bmp file. You must enter a file name and add the extension ".bmp" to ensure the file is saved as an image file. If your system has QuickTime or you are using a Mac, you will have other file type options such as gif and jpg.

Paste Chart

If using Windows, you may need to use Paste Special/Bitmap to paste the chart into other software such as Microsoft Word. Mac users should only need to use Paste to insert the chart into a document.

Variable Limits

Some analyses only use one variable (e.g., column and pie charts), while others use two variables (scatterplot). You cannot run multiple charts, so AcaStat will limit the number of variables you can select. You must remove a variable from the listbox before replacing it with another variable.

Editing Charts

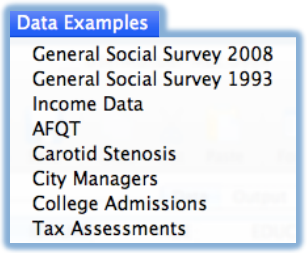
Most features of the charts are automatic based on the data and the labeling of variables and values in your data file. However, you can change some settings by clicking the Edit Chart button to view chart options.

Options include changing the chart title, axis labels, and the chart background. You may also set decimal values and minimum and maximum values for the Y and X axis, but this is currently only available for scatterplots.

Click OK to rebuild the chart using the settings you provide. Click Cancel to close without making the changes. The chart procedure will return to the default settings when you run a new chart.

Other Menu Items and AcaStat Features

Data Examples Menu



Several practice data files are available for use with AcaStat. The data files include variable and value formatting. The files can be saved for manipulation and use as instructional material.

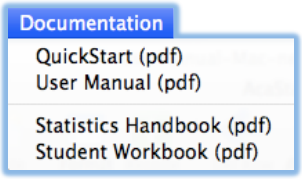
The following data sets are based on actual research surveys.

General Social Surveys 2008 and 1993
AFQT
City Managers
College Admissions

The following data sets are fictional.

Carotid Stenosis
Tax Assessments
Income Data

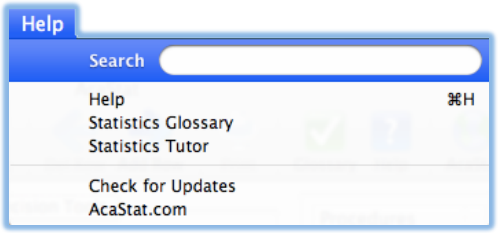
Documentation



Use this pull-down menu to open application documentation and statistics instructional material.

Help Menu

Help



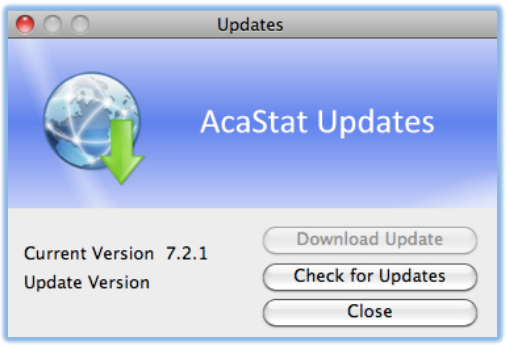
Click Help to open the help module. The contents can be sorted or searched by topic.

Statistics Glossary

The glossary contains over 150 terms and definitions. The Print button will print the item being viewed. The search button will find matching records and sort the listbox so all topics matching the search are listed at the top in bold. Pressing the Enter key after entering a search term will also initiate the search. Press Reset before conducting another search.

Check for Updates

The Help menu contains an option to check for updates. AcaStat connects to the Internet and automatically compares the installed version with the version of the update. If a newer update is available, the appropriate installation file is downloaded to the desktop of your computer.



The update option is only available for licensed users.

Applied Statistics Handbook

The Applied Statistics Handbook contains formulas, definitions, and examples along with annotated output from analytical software. It is designed to be a quick reference resource (not a complete statistical text).

Click this menu item to access the Handbook pdf. If the program is not a licensed version, this option will not be operational. All licensed users also have the pdf file installed on their computer during the AcaStat installation process. Mac users will find the Handbook under Applications in the "AcaStat Library" folder. Windows users will find the Handbook in the AcaStat Folder in the Program Files directory.

File Name: Statbook.pdf
Availability: AcaStat Licensed Edition

Student Workbook

The Student Workbook contains over 20 lessons designed to introduce basic statistical concepts while also demonstrating the functionality of AcaStat.

Click this menu item to access the Workbook pdf. All users have the pdf file installed on their computer during the AcaStat installation process. Mac users can find the Workbook under Applications in the "AcaStat Library" folder. Windows users will find the Student Workbook in the AcaStat Folder in the Program Files directory.

File Name: Workbook.pdf
Availability: AcaStat CD, Website http://www.acastat.com/pdfdocs.htm

Review of Features

AcaStat is designed to provide students, teachers, researchers, and managers with an inexpensive and easy to use tool for data exploration. It includes three data analysis modules (Data Grid, SumStats, Decision Tools) and three support modules containing analysis output (Output Viewer tab), statistical terms and definitions (Glossary module), and a tutorial. Special pricing and site licenses are available for instructors adopting AcaStat for course work. Printable manuals include a Student Workbook with over 20 lessons, an Applied Statistics Handbook, and an AcaStat User Manual.

Data Spreadsheet

Analyze raw data with this spreadsheet design. Produces basic descriptive and inferential statistics, and creates crosstabulations. Includes the features listed below.

Feature	Details
Spreadsheet	create data files, maximum edit capacity 60 columns
Import Data	import comma and tab delimited text files, paste tables from other software
Data Formatting	create variable and value labels, set missing values
Variable Recode	compute, combine, and recode variables
Paste Data	copy and paste data from spreadsheets, documents, and database tables
Export Data	export comma and tab delimited text files
Output	edit output, save as text file, print, copy, paste
Statistics	
Frequencies	frequencies (count, percent, total)
Descriptives	mean, median, variance, standard deviation
t-test of means	
One-way ANOVA	summary statistics and Bonferroni post hoc analysis
Chi-square	Cramer's V, Pearson C, Lambda, Kappa, Fisher's, odds ratio
Pearson correlation	correlation matrix and scatterplot
Simple regression	OLS
Multiple regression	OLS and Logistic
Sampling	Random sample and repeated random sampling
Explore Means	
Diagnostic Accuracy	Sensitivity, specificity, likelihood ratios, predictive values
Appraisal	Ratios and summary statistics

SumStats

Analyze summary statistics (means, proportions, counts) to develop confidence intervals and conduct basic significance tests. Includes the features listed below. SumStats provides summary instructions and interpretations of output.

Feature	Details
Statistical Tests	one and two sample t-test of means one-way ANOVA (6 sample) one and two sample z-test of proportions chi-square (6x6 table)
Raw Data Analysis	Descriptive statistics Correlation OLS Simple and Multiple Regression
Margin of Error-Mean	90%-95%-99% confidence intervals, margin of error, upper limits, lower limits
Margin of Error-Proportions	90%-95%-99% confidence intervals, margin of error, upper limits, lower limits
Standardized Scores	z-score, graphic representation on normal distribution
Diagnostic Accuracy	sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, negative likelihood ratio

System Requirements (recommended minimum)

Mac OS X v. 10.2 or later
Windows 98 or later
Download file size 10 MB; Installed size 30 MB

AcaStat Software
43584 Merchant Mill Terrace
Leesburg, VA 20176 USA
support@acastat.com
http://www.acastat.com

Purchase Information

Prices are subject to change. Special orders, quantity discounts, and site licenses are available. Questions? Contact AcaStat Software at info@acastat.com.

On line Orders (Credit Card)

Visit http://www.acastat.com for credit card orders, current prices, and site license purchases and renewals.

Mail orders (check, money-order, or purchase order/invoice)

Use the Mail Order Form

AcaStat Software Mail Order Form

Name:_____ E-Mail: _____

Organization/School: _____

Street: _____ Phone: _____

City: _____ State/Province: _____

Country: _____ Postal Code: _____

Download - Retail $29.95 ea ____ x 29.95 = _____

Download – Academic $19.95 ea ____ x 19.95 = _____

Add Mini CD $10.00 ea ____ x 10.00 = _____

Subtotal = _____

Virginia residents add 5% sales tax (subtotal x .05) = _____

Total = _____

Make Check or Money Order Payable to:

AcaStat Software
43584 Merchant Mill Terrace
Leesburg, VA 20176 USA